PRECISION MEASUREMENT OF THE HIGGS BOSON MASS AND SEARCH FOR DILEPTON MASS RESONANCES IN $H \rightarrow 4\ell$ DECAYS USING THE CMS DETECTOR AT THE LHC

By

JAKE ROSENZWEIG

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2022

I dedicate this to Jacob Myhre.

ACKNOWLEDGEMENTS

Without so many two- and four-legged blessings along the way, I could never have made it to this point in my academic career. Thus, I give my infinite thanks to the following folks.

To my high energy physics mentors, Professors Andrey Korytov and Guenkah Mitselmakher, for granting me this one-of-a-kind opportunity to do real *science* at CERN.

To my wife, Suzanne Rosenzweig, for showing me that dreams do come true. To my mother and father, Vicki and John, who always reassured me that I could achieve anything I put my mind to. Sleep peacefully, Mom. To my siblings, Alex, Ryan, Devin, Jace, and Claudia who frequently and gently reminded me that there was life outside of grad school.

Aunt Rach, Uncle Yuri,

To my mentor, Sheldon Friedman, and his wife, Rita Friedman (Rosenzweig), who chose to invest in my success at a young age. I have only made it this far thanks to your undying encouragement, love, and optimism. To Sheldon's best friend, Dr. Bernard Khoury, whose reputation has helped pave my road.

To Dr. Filippo Errico for his Dr. Lucien Lo To Dr. Noah Steinberg To Darin Acosta, for spending hours of discussion on every t To the gentle gents who introduced me to the world of CMS, Brendan Regnery and Bhargav Joshi.

To my comrades for showing me what it takes to survive the core courses, Dr. Atool Divakarla, Dr. Brien O'Brendan, Dr. Donyell Guerrero, and Dr. Vladinar Martinez.

To the many students who tagged along in our "CMS Office Hours": Sean Kent, Jeremiah Anglin, Cris Caballeros, Ari Gonzalez, Evan Koenig, Nik Menendez, Neha Rawal, John Rötter. And to the many students who let us practice our spiels:

To my mentee, Matthew Dittrich, for

To my Polish roommates in Saint-Genis-Pouilly for showing me what home away from home feels like.

To the many moms who generously gave unconditional support during the darkest times and unconditional love during the brightest times: Silet Wiley, Margaret Sherrill, Dawn Hood, Cyndi Reilly-Rogers.

To my childhood best friends: Jish, Willis, Shane, Zac, Duck, and Marcus for their constant

clever competition which has shaped me into the determined man I am today.

Big Tree:

And finally to Existence itself for this unpredictable, unbelievable blip of an experience called life.

TABLE OF CONTENTS

page

LIST OF TABLES

LIST OF FIGURES

Figures page

CHAPTER 1
INTRODUCTION

The universe, while overwhelmingly vast, is comprised of a curiously small number of elementary particles. These particles and their strong, weak, and electromagnetic interactions with each other are accurately described by the Standard Model (SM). A major shortcoming of the SM was its inability to predict the masses of these particles.

This dissertation presents a precision measurement of the Higgs boson mass and using LHC proton-proton collision data from Run 2 data set from

The SM was not able to predict the masses of these particles until 1964 when the Brout-Englert-Higgs mechanism suggested that It wasn't until 1964 that the Brout-Englert-Higgs mechanism gave a self-consistent way to : by breaking the electroweak gauge symmetry of the vacuum would give rise to non-zero masses of the weak gauge bosons. This would yield a secondary effect too: there should exist a fundamental scalar boson which is the quantum of the so-called "Higgs field". On July 4th, 2012, this Higgs boson was discovered.

At first glance, the universe appears to be an overwhelmingly vast and complicated place. However upon closer inspection, it is comprised of only a few different kinds of fundamental particles. Particle physics has given rise to the Standard Model (SM) which mathematically describes these constituents and their interactions with each other.

The Standard Model (SM) is an impressively accurate mathematical theory which describes the fundamental particles of the universe and the rules for their possible interactions. Problematically though, the SM predicts that all particles are massless.

Get to the Higgs boson.

Why is it important? Knowing the mass of the Higgs boson

CHAPTER 2
HIGGS BOSON MASS MEASUREMENT IN THE $H \rightarrow ZZ^{*} \rightarrow 4\ell$ CHANNEL

2.1 Introduction

The Higgs boson was discovered in 2012 by the CMS and ATLAS collaborations. This was a momentous achievement in particle physics because the existence of the Higgs boson was required to complete the SM. In fact, it is sometimes referred to as the "missing puzzle piece" of the SM. The Higgs boson is one of a kind: it is the only fundamental scalar particle ever discovered so far. The unique boson could be a portal to new physics (*beyond Standard Model physics*, BSM), e.g., by decaying into BSM low-mass dilepton mass resonances (Chapter 3). In order to be certain that the recently discovered Higgs boson is truly the same one predicted by the SM, it is necessary to compare its measured properties to the predicted values.

Some properties of the Higgs boson can be predicted by the SM, like - There are many results on Higgs properties: spin, charge, decay processes, lifetime, mass. - The last of these is of particular importance: depending on mH and mtop, the stability of the Universe.

- ALL previous mass measurements: - Run 1: - H->2gamma VALUE - H->ZZ->4L VALUE - Run 2: - H->2gamma VALUE - (2016) H->ZZ->4L VALUE - H->bb - H->mumu - H->WW

- Why this thesis is important: - This thesis describes the methodology and results of the best precision measurement of mH to date by using the hZZ4l decay and Full Run 2 data set from CMS. - Run 2 provides more data -> more precision on measurements of Higgs properties. - In addition to more HZZ4l events, this analysis provides new techniques, specifically the VX constraint. - Predict mH for Run 3, will start soon summer 2022 and provide an approximate 300? /fb of L int. - In 2026(?), HLLHC provides even more data. ref snowmass paper.

This chapter is structured as follows:

- General overview of the ingredients of the Higgs boson mass measurement (Section 2.2).
- Data sets, simulated samples, triggers (Section **??**).
- Event reconstruction and selection (Section 2.4).
- Background Estimation (Irred. and Reduc. Backgrounds).
- Signal Modeling: kinematic Discriminant, per-event mass uncertainties, VXBS constraint, reference to ad hoc studies in appendix.

- Systematic Uncertainties.
- Results.
- Summary.

SEEMS TO BE A GOOD INTRO. Should it be the intro for the entire thesis?

2.2 Analysis Overview

The first step to performing a precision measurement of the Higgs boson mass is to "observe" many Higgs bosons. Since the Higgs boson has a *very* short mean lifetime of only 1×10^{-23} s [1],

- Want to measure the Higgs boson mass (mH), so need Higgs bosons. - Sift through CMS data for $\mathrm{H} \to \mathrm{ZZ}^* \to 4\ell$ events (the *signal* process) because S B ratio is huge: 2. - However $\mathrm{H} \to \mathrm{ZZ}^* \to 4\ell$ is rare: Although the LHC , the Higgs boson is produced in only 1 out of every billion pp collisions. Even if H is produced, it will decay into two Z only a small percentage of the time (2.3%). This percentage is typically expressed as a fraction, called the *branching fraction* or *branching ratio* ($\mathcal{B}$). there is a small probability of only that it will decay into two Z bosons (2.3%). Those Z bosons then have only a small probability so the boson itself will never live long enough to interact directly with the CMS Detector. Therefore, the Higgs boson can only be detected by the daughter particles into which it decays. - By collecting events with the 4ℓ final state, we are likely to find signal events. - It's not just the signal process which produces 4ℓ: background also makes 4ℓ (Section FIXME). - Before analyzing the data, however, it is important to make predictions using simulated samples (Section FIXME). - In order to sort signal from background, use simulated samples

- Form objects from data. - Use objects and conservation of momentum to rebuild parent particles. - The Z boson has a precisely measured mass of TODO a neutral particle, so the two leptons into which it decays should combine to Group two leptons together, - Form two different pairs of opposite-sign, same-flavor (OSSF) leptons - If it appears that the to select specific hzz4l events (*event selection*). -

2.3 Data Sets, Simulated Samples, and Triggers

2.3.1 Data Sets

2.3.2 Simulated Samples

2.3.3 Triggers

2.4 Event Reconstruction and Selection

2.4.1 Event Reconstruction

2.4.2 Event Selection

2.4.2.1 ZZ Candidate Selection

2.5 Background estimation

Measurement of the Higgs boson mass requires the accurate modelling of the total event yield in the signal region (SR), into which events can be categorized as either *signal* or *background*. These background events pass the signal event selection (Section 2.4.2.1) and thus spoil the purity of the signal events. This introduces further uncertainty into the final Higgs boson mass measurement. Therefore, it is a priority to both reduce and predict the expected number of background events, which can be split into *irreducible backgrounds* (IB) and *reducible background* (RB) processes.

2.5.1 Irreducible background

IB processes produce two Z bosons and each Z subsequently decays into two prompt leptons (leptons that emerge directly from the primary vertex). This reliably produces a 4ℓ final state, whose prompt leptons typically get reconstructed as four leptons that pass tight selection (*PTS leptons*), as defined in Sections **??** and **??**. The IB event then looks indistinguishable from the 4 PTS lepton of the *signal* process and cannot be reduced; Thus, throwing away IB process could mean throwing away a signal event. Since IB event cannot be reduced, it's called irreducible background. The two IBs for the HZZ4L analysis are:

- $\mathrm{gg} \rightarrow \mathrm{ZZ} \rightarrow 4\ell$ (gluon-gluon fusion),
- $\mathrm{q\bar{q}/gg} \rightarrow \mathrm{ZZ} \rightarrow 4\ell$ (quark-antiquark annihilation).

2.5.2 Reducible background

While irreducible background processes produce 4 prompt leptons, RB processes produce varying numbers of prompt and nonprompt leptons. Nonprompt leptons emerge from 3 main sources:

- misidentifying light-flavor hadrons (e.g., $\pi^{\pm}$) as leptons,
- heavy-flavor hadrons that decay mid-flight into leptons,
- and the asymmetric conversion of photons into electrons.

This analysis is tailored to efficiently reconstruct prompt leptons as those which pass tight selection (*PTS leptons*), whereas nonprompt leptons typically fail tight selection (*FTS leptons*). Tight and loose lepton selections are defined in Sections **??** and **??**. Ideally, the SR would contain only the events with 4 prompt leptons, which would always be tagged as 4 PTS leptons. However, sometimes a truly nonprompt lepton is misidentified as a PTS lepton (sometimes called a *fake lepton*), depending on the kinematical properties of the lepton. This misidentification rate (f, sometimes called the *fake rate*) is due to imperfect detector performance, inefficiencies in reconstruction, and the specific choice of lepton selections used in the analysis.

If an event produces prompt and nonprompt leptons but is reconstructed as 4 PTS leptons (a 4P event), then it contaminates the SR. These processes constitute the RB processes (sometimes called "$\mathrm{Z}+\mathrm{X}$"). Examples of RB processes for this analysis include:

- $\mathrm{Z}+\mathrm{jets}$ (yields 2 prompt leptons)
- $\mathrm{t\bar{t}}+\mathrm{jets}$ (yields 2 prompt leptons)
- $\mathrm{WZ}+\mathrm{jets}$ (yields 3 prompt leptons)
- $\mathrm{Z}(\mathrm{Z}/\gamma^{*})+\mathrm{jets}$ (yields 4 prompt leptons).

The careful estimation of these RB contributions to the 4P region is necessary for the precise measurement of the Higgs boson mass.

2.5.2.1 OS Method

The goal of the OS Method is to estimate the number of opposite-sign same-flavor (OSSF) 4ℓ events produced by RB process that "contaminate" the 4P region $\left(N_{\mathrm{4P}}^{\mathrm{RB}}\right)$, given by Eq. 2-13. However, the typical approach of using simulated samples does not model RB well, since RB processes (e.g., $\mathrm{Z}+\mathrm{jets}$) rely on higher-order effects like jet modelling which are not yet accurately simulated. Instead, a data-driven approach is used.

The logic of the OS Method is to study events in data that are similar to, but not exactly the same as, those found in the 4P region. Thus, events in data are sorted into 2 control regions (CRs), both of which are orthogonal to the 4P region and to each other:

- the *3P1F* CR (built from 3 PTS and 1 FTS leptons)
- the *2P2F* CR (built from 2 PTS and 2 FTS leptons).

The event selection for the 2P2F and 3P1F CRs is almost identical to that of the SR (Section 2.4.2.1), except that the FTS lepton(s) are required to build the Z_2. The events that contribute mostly to the 2P2F(3P1F) CR are those that produce 2(3) prompt and 2(1) nonprompt leptons and are called *2pr*(*3pr*) events. Similarly, the events that contribute mostly to the 4P SR are those that produce 4 prompt leptons and are called *4pr* events.

The final formula for $N_{\mathrm{4P}}^{\mathrm{RB}}$ is obtained by first supposing that event k is a 2pr event and contributes to the 2P2F CR an event weight of $w^k_{\mathrm{2pr}\to\mathrm{2P2F}}$. This weight is built from the product of analysis weights (pileup, L1 pre-firing, etc., whose product is $\hat{w}^k$) and the reconstruction efficiencies (ϵ) of each lepton (ℓ_n):

$$w^k_{\mathrm{2pr}\to\mathrm{2P2F}} = \hat{w}^k \cdot \epsilon_{\mathrm{P}}^{\mathrm{pr}}(\ell_1^k) \cdot \epsilon_{\mathrm{P}}^{\mathrm{pr}}(\ell_2^k) \cdot \epsilon_{\mathrm{F}}^{\mathrm{np}}(\ell_3^k) \cdot \epsilon_{\mathrm{F}}^{\mathrm{np}}(\ell_4^k), \tag{2-1}$$

where the superscript of ϵ indicates the lepton promptness (pr = prompt, np = nonprompt), the subscript indicates the lepton tightness status (P = PTS, F = FTS). To simplify the equations that follow, $\hat{w}^k$ is set to unity. If the reconstruction efficiencies of a particular category are the same for all j leptons across all k events $\left(\text{e.g., } \epsilon_{\mathrm{P}}^{\mathrm{pr}}(\ell_j^k) \equiv \epsilon_{\mathrm{P}}^{\mathrm{pr}}\right)$, then Eq. 2-1 reduces to:

$$w^k_{\mathrm{2pr}\to\mathrm{2P2F}} = \left(\epsilon_{\mathrm{P}}^{\mathrm{pr}}\right)^2 \left(\epsilon_{\mathrm{F}}^{\mathrm{np}}\right)^2. \tag{2-2}$$

Although a 2pr event mostly contributes to the 2P2F CR, a nonprompt lepton may be misidentified as a PTS lepton, depending on $\epsilon_{\mathrm{P}}^{\mathrm{np}}$. Such an event would then fall into the 3P1F CR and, allowing for only one nonprompt PTS lepton at a time, contributes an effective weight of

$$\begin{aligned} w^k_{\mathrm{2pr}\to\mathrm{3P1F}} &= \left(\epsilon_{\mathrm{P}}^{\mathrm{pr}}\right)^2 \left[\epsilon_{\mathrm{P}}^{\mathrm{np}}(\ell_1^k)\cdot\epsilon_{\mathrm{F}}^{\mathrm{np}}(\ell_2^k) + \epsilon_{\mathrm{F}}^{\mathrm{np}}(\ell_1^k)\cdot\epsilon_{\mathrm{P}}^{\mathrm{np}}(\ell_2^k)\right] \\ &= \left(\epsilon_{\mathrm{P}}^{\mathrm{pr}}\right)^2 \left[2\epsilon_{\mathrm{P}}^{\mathrm{np}}\epsilon_{\mathrm{F}}^{\mathrm{np}}\right]. \end{aligned} \tag{2-3}$$

Using the fact that a (non)prompt lepton is exclusively either PTS or FTS $\left(\epsilon_{\mathrm{P}}^{\mathrm{pr(np)}} + \epsilon_{\mathrm{F}}^{\mathrm{pr(np)}} = 1\right)$, while recognizing that $\epsilon_{\mathrm{P}}^{\mathrm{np}} \equiv f$ (Section 2.5.2) and defining $\epsilon_{\mathrm{P}}^{\mathrm{pr}} \equiv \epsilon$, allows Eq. 2-3 to be written as

$$w^k_{\mathrm{2pr}\to\mathrm{3P1F}} = 2\epsilon^2 f(1-f). \tag{2-4}$$

The prediction of Eq. 2-4 can be seen in Figures 2-5–2-7 for 2016–2018 UL data.

Even more rarely, both nonprompt leptons from a 2pr event may be misidentified as PTS leptons. In this case, event k contributes to the 4P region an effective weight of

$$w^k_{\mathrm{2pr}\to\mathrm{4P}} = \epsilon^2 f^2, \tag{2-5}$$

where it is assumed that both leptons have the same misidentification rate. Similar equations can be derived for the contributions of a 3pr event to the 3P1F CR and to the 4P region:

$$w^k_{\mathrm{3pr}\to\mathrm{3P1F}} = \epsilon^3(1-f) \tag{2-6}$$

$$w^k_{\mathrm{3pr}\to\mathrm{4P}} = \epsilon^3 f. \tag{2-7}$$

Since a 3pr event needs only 1 nonprompt PTS lepton to be included in the 4P region (therefore, carrying only 1 factor of f), a 3pr event tends to contribute more weight to 4P than does a 2pr event (which carries f^2).

If the total number of 2pr, 3pr, and 4pr events is X_{2pr}, X_{3pr}, and X_{4pr}, respectively, then

Figure 2-1 shows how the weight of a single event, derived in the previous equations (and others forthcoming), from each category contributes to the final yield of each CR ($N_{\mathrm{2P2F}}, N_{\mathrm{3P1F}}$) and of the SR ($N_{\mathrm{4P}}$). It is then straightforward to evaluate the expected number of RB 4P events:

$$\begin{aligned} N_{\mathrm{4P}}^{\mathrm{RB}} &= \sum_{k=1}^{X_{\mathrm{2pr}}} w_{\mathrm{2pr}\to\mathrm{4P}}^{k} + \sum_{m=1}^{X_{\mathrm{3pr}}} w_{\mathrm{3pr}\to\mathrm{4P}}^{m} \\ &= \sum_{k=1}^{X_{\mathrm{2pr}}} \epsilon^2 f^2 + \sum_{m=1}^{X_{\mathrm{3pr}}} \epsilon^3 f \\ &= f^2\epsilon^2 X_{\mathrm{2pr}} + f\epsilon^3 X_{\mathrm{3pr}}, \end{aligned} \tag{2-8}$$

where only the quantities f and ϵ are known, so X_{2pr} and X_{3pr} must be estimated. This is achieved by relating X_{2pr} to N_{2P2F} using Eq. 2-2 across all 2pr events:

$$\begin{aligned} N_{\mathrm{2P2F}} &= \sum_{k=1}^{X_{\mathrm{2pr}}} w_{\mathrm{2pr}\to\mathrm{2P2F}}^{k} \\ &= \sum_{k=1}^{X_{\mathrm{2pr}}} \epsilon^2 (1-f)^2 \\ &= (1-f)^2 \epsilon^2 X_{\mathrm{2pr}} \end{aligned} \tag{2-9}$$

The strategy to relate X_{3pr} to N_{3P1F} is not as straightforward as relating X_{2pr} to N_{2P2F}, since two other sources also contribute to the 3P1F CR (as shown in Figure 2-1):

- A 2pr RB process can yield one nonprompt PTS lepton (via Eq. 2-4).
- A 4-prompt IB process ($\mathrm{q\bar{q}/gg} \to \mathrm{ZZ} \to 4\ell$, “ZZ”) can yield one FTS lepton.

The second of these is well estimated from simulation, since ZZ produces 4 prompt leptons. If the total number of simulated ZZ events is $X_{\mathrm{4pr}}^{\mathrm{ZZ}}$, then event k belonging to these events contributes to the 3P1F CR an effective weight of

$$w_{\mathrm{4pr,ZZ}\to\mathrm{3P1F}}^{k} = 4\epsilon^3 \epsilon_{\mathrm{F}}^{\mathrm{pr}}, \tag{2-10}$$

which accounts for any of the 4 prompt leptons to be reconstructed as a FTS lepton while the

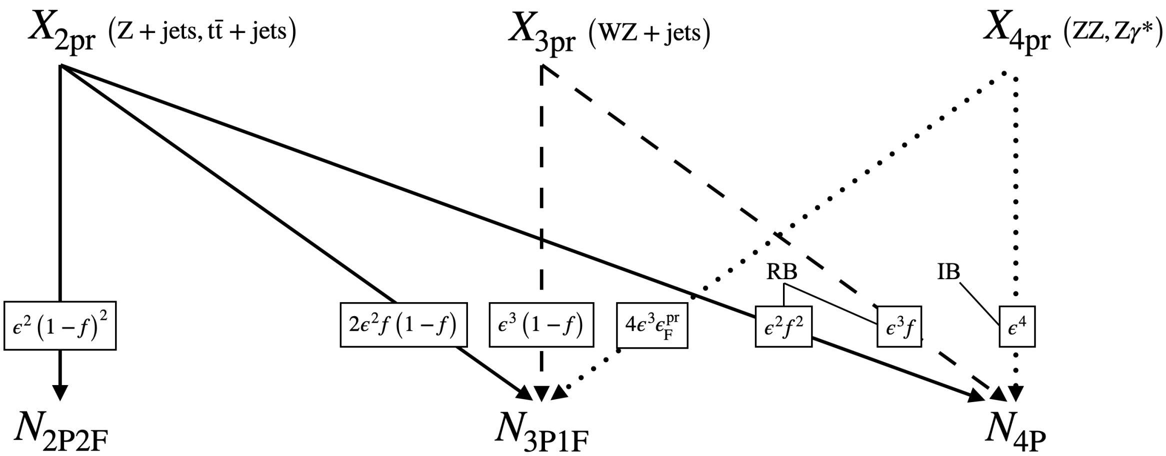

Figure 2-1. Contributions of per-event weights (boxed values) of various n-prompt-lepton processes ($X_{n\text{pr}}$, in parentheses) to the total numbers of events in the observed control regions ($N_{2\text{P2F}}, N_{3\text{P1F}}$) and signal region ($N_{4\text{P}}$). The labels "IB" and "RB" indicate those contributions which comprise the irreducible and reducible backgrounds, respectively.

others are PTS leptons. Incorporating Eqs. 2-4, 2-6, and 2-10 for all events relates $X_{3\text{pr}}$ to $N_{3\text{P1F}}$:

$$\begin{aligned}
N_{3\text{P1F}} &= \sum_{j=1}^{X_{3\text{pr}}} w^j_{3\text{pr}\to 3\text{P1F}} + \sum_{k=1}^{X_{2\text{pr}}} w^k_{2\text{pr}\to 3\text{P1F}} + \sum_{m=1}^{X^{\text{ZZ}}_{4\text{pr}}} w^m_{4\text{pr},\text{ZZ}\to 3\text{P1F}} \\
&= \sum_{j=1}^{X_{3\text{pr}}} \epsilon^3(1-f) + \sum_{k=1}^{X_{2\text{pr}}} 2\epsilon^2 f(1-f) + \sum_{m=1}^{X^{\text{ZZ}}_{4\text{pr}}} 4\epsilon^3\epsilon^{\text{pr}}_{\text{F}} \\
&= (1-f)\epsilon^3 X_{3\text{pr}} + 2f(1-f)\epsilon^2 X_{2\text{pr}} + 4\epsilon^3\epsilon^{\text{pr}}_{\text{F}} X^{\text{ZZ}}_{4\text{pr}} \\
&= (1-f)\epsilon^3 X_{3\text{pr}} + 2f(1-f)\epsilon^2 X_{2\text{pr}} + N^{\text{ZZ}}_{3\text{P1F}},
\end{aligned} \tag{2-11}$$

where $N^{\text{ZZ}}_{3\text{P1F}}$ is simply the raw (integer) number of ZZ events that pass 3P1F selections, obtained directly from simulation.

At this point, $N^{\text{RB}}_{4\text{P}}$ can be isolated by combining Eqs. 2-8, 2-9, and 2-11:

$$N^{\text{RB}}_{4\text{P}} = \left(\frac{f}{1-f}\right) N^{\text{Data}}_{3\text{P1F}} - \left(\frac{f}{1-f}\right)^2 N^{\text{Data}}_{2\text{P2F}} - \left(\frac{f}{1-f}\right) N^{\text{ZZ}}_{3\text{P1F}}. \tag{2-12}$$

Thus, Eq. 2-12 estimates the RB contribution to the 4P region by using a single lepton misidentification rate to reweight the raw yields of events found in the 3P1F and 2P2F CRs of data, and also the 3P1F CR of ZZ.

It should be mentioned that the above formula was derived assuming that the per event analysis weights were set to unity $\left(\hat{w}^k = 1\right)$ before being scaled by the misidentification rates. It was also assumed that f is a constant, for all nonprompt leptons misidentified as PTS leptons across all events, which is not the case as is shown in Figure 2-8. Thus, an extension of Eq. 2-12 can be formed by assigning misidentification rates that depend on the kinematical variables per lepton, by restoring the analysis weights per event, and by summing over the total number of raw yields per CR:

$$N_{\text{4P}}^{\text{RB}} = \sum_{i=1}^{N_{\text{3P1F}}^{\text{Data}}} \hat{w}^i \frac{f_i}{1-f_i} - \sum_{j=1}^{N_{\text{2P2F}}^{\text{Data}}} \hat{w}^j \frac{f_{1,j}}{1-f_{1,j}} \frac{f_{2,j}}{1-f_{2,j}} - \sum_{k=1}^{N_{\text{3P1F}}^{\text{ZZ}}} \hat{w}_{\text{ZZ}}^k \frac{f_k}{1-f_k}, \tag{2-13}$$

where $f_{1,j}$ and $f_{2,j}$ are the misidentification rates of the first and second FTS leptons, respectively, found in the j^{th} 2P2F event and $\hat{w}_{\text{ZZ}}^k$ accounts for the differential QCD and electroweak k-factors defined in Section 2.5.1.

2.5.2.2 Lepton misidentification rate measurement

As mentioned in Section 2.5.2, the lepton misidentification rate (f) is the probability that a nonprompt lepton will pass tight selections (PTS). The value f is a function of the flavor $(\ell = \text{e}, \mu)$, p_{T}, and η of a lepton. The misidentification rate is calculated by simply counting the number of nonprompt PTS leptons $\left(N_{\text{P}}^{\text{np}}\right)$ that enter a particular $\ell, p_{\text{T}}, |\eta|$ bin compared to the total number of loose probe leptons $\left(N_{\text{L}}^{\text{np}}\right)$ in the same bin:

$$f\left(\ell, |\eta|, p_{\text{T}}\right) = \frac{N_{\text{P}}^{\text{np}}}{N_{\text{L}}^{\text{np}}}. \tag{2-14}$$

The p_{T}^{e} bin edges are [5–10–20–30–40–50–80] GeV and the p_{T}^{μ} bin edges are [5–7–10–20–30–40–50–80] GeV. The nonprompt leptons used to measure f are taken from events in data with a signature like that of $\text{Z} + \ell_{\text{L}}$, where Z is a Z boson and ℓ_{L} is a loose lepton. By construction, this region of events is orthogonal to the 2P2F, 3P1F, and 4P regions, and provides a clean source of ℓ_{L}. The loose lepton, whose selection is defined in Sections **??** and **??**, is also called the *probe* lepton. The probe lepton is either a PTS or FTS lepton and is counted

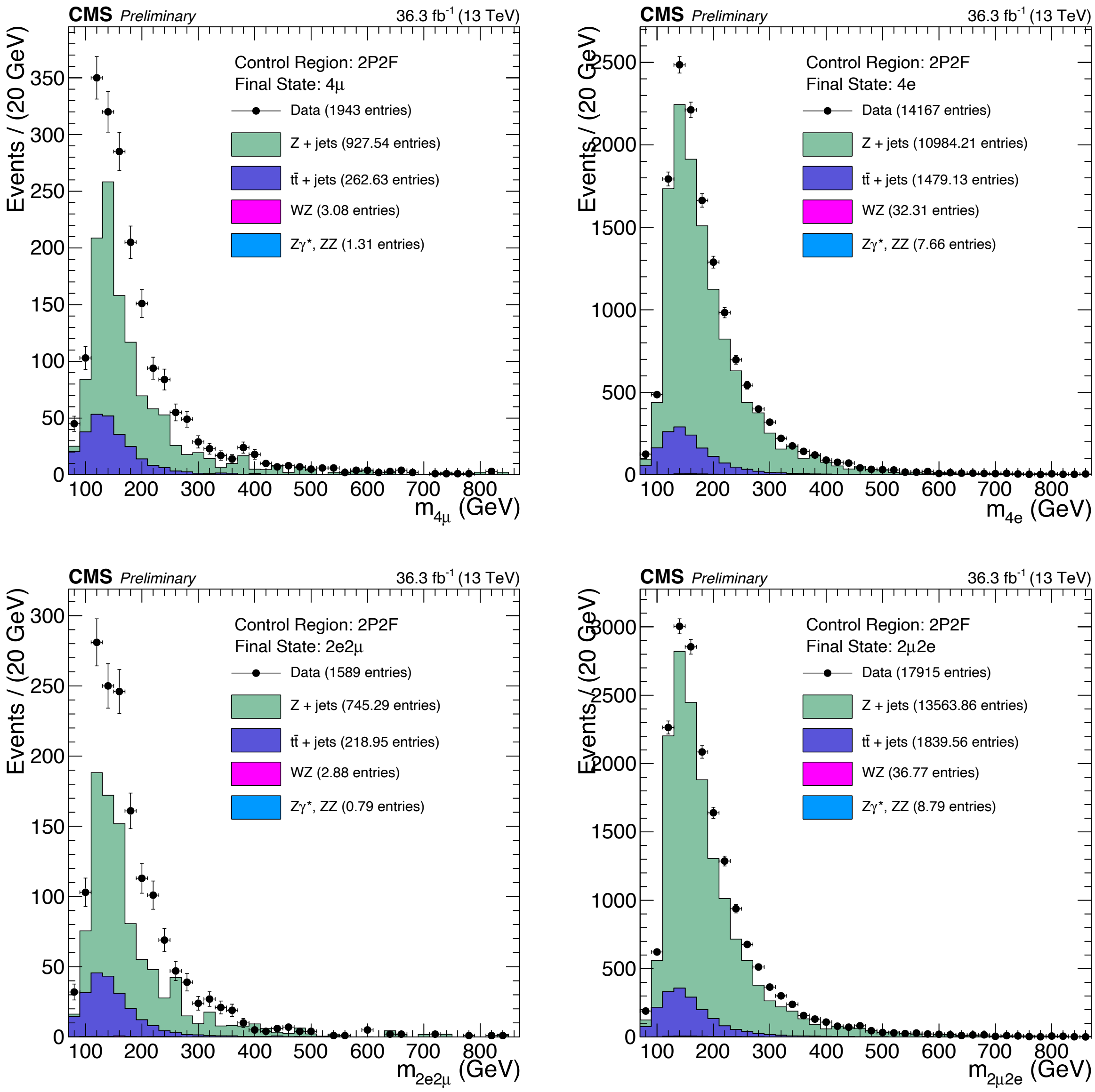

Figure 2-2. Events from 2016 UL data that pass 2P2F CR selections (black markers) are compared to the stacked 2P2F distributions of simulated samples (Z + jets, $t\bar{t}$ + jets, WZ, Z/γ^*, ZZ). The results are split into the 4ℓ final states: 4μ (top left), 4e (top right), $2e2\mu$ (bottom left), $2\mu 2e$ (bottom right).

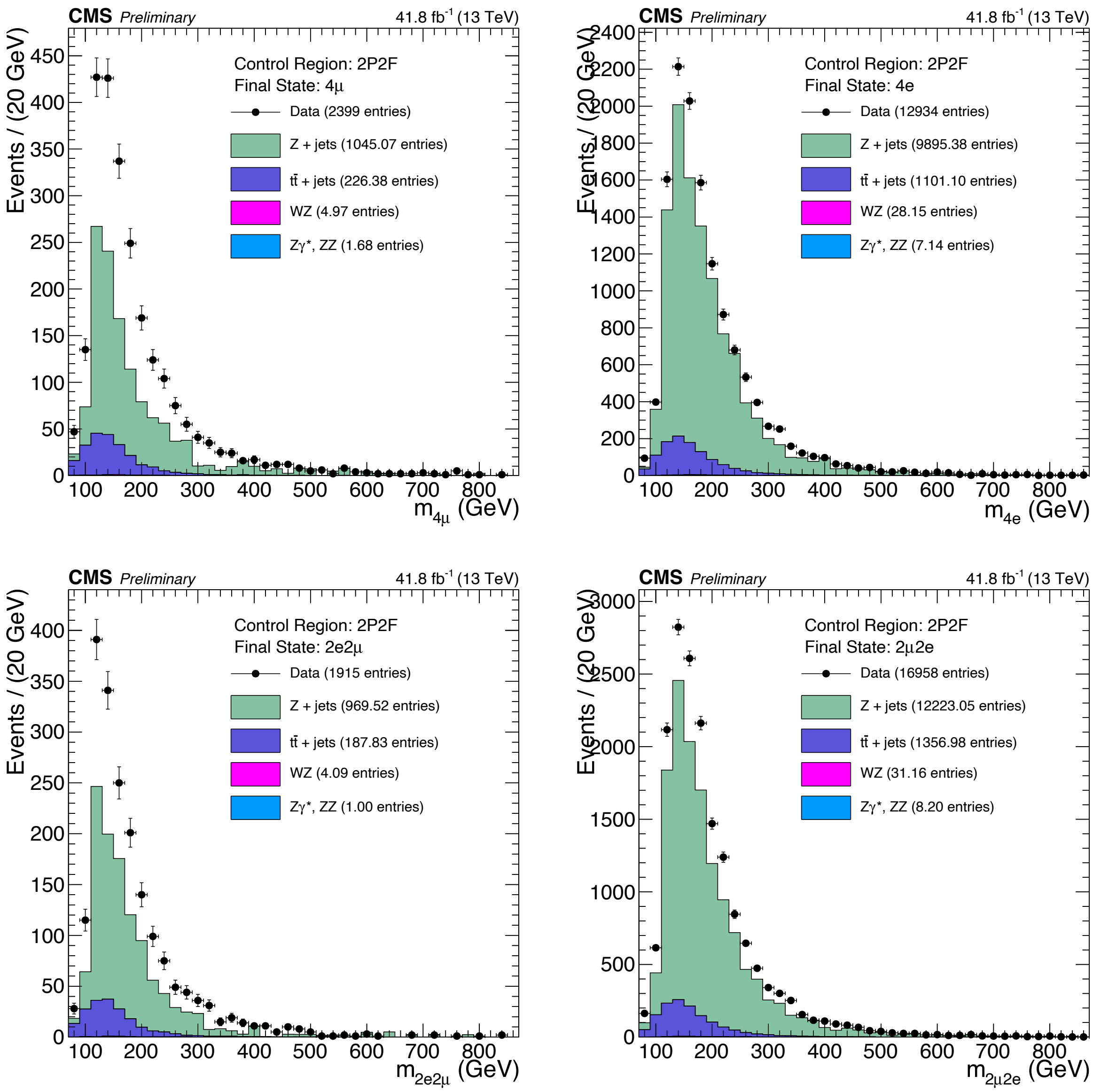

Figure 2-3. Events from 2017 UL data that pass 2P2F CR selections (black markers) are compared to the stacked 2P2F distributions of simulated samples (Z + jets, $t\bar{t}$ + jets, WZ, Z/γ^*, ZZ). The results are split into the 4ℓ final states: 4μ (top left), 4e (top right), 2e2μ (bottom left), 2μ2e (bottom right).

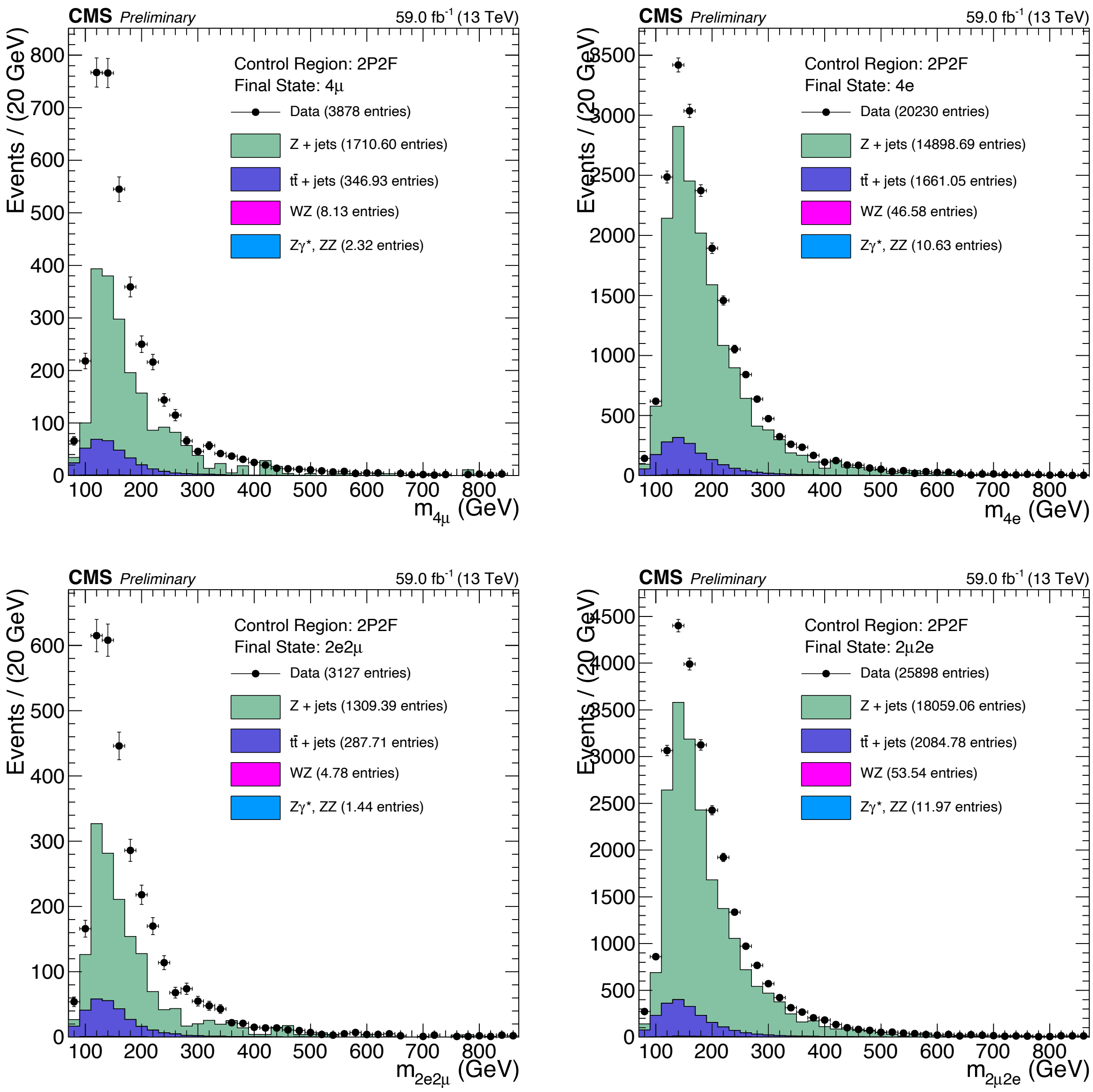

Figure 2-4. Events from 2018 UL data that pass 2P2F CR selections (black markers) are compared to the stacked 2P2F distributions of simulated samples ($Z+\text{jets}$, $t\bar{t}+\text{jets}$, WZ, Z/γ^*, ZZ). The results are split into the 4ℓ final states: 4μ (top left), 4e (top right), $2e2\mu$ (bottom left), $2\mu2e$ (bottom right).

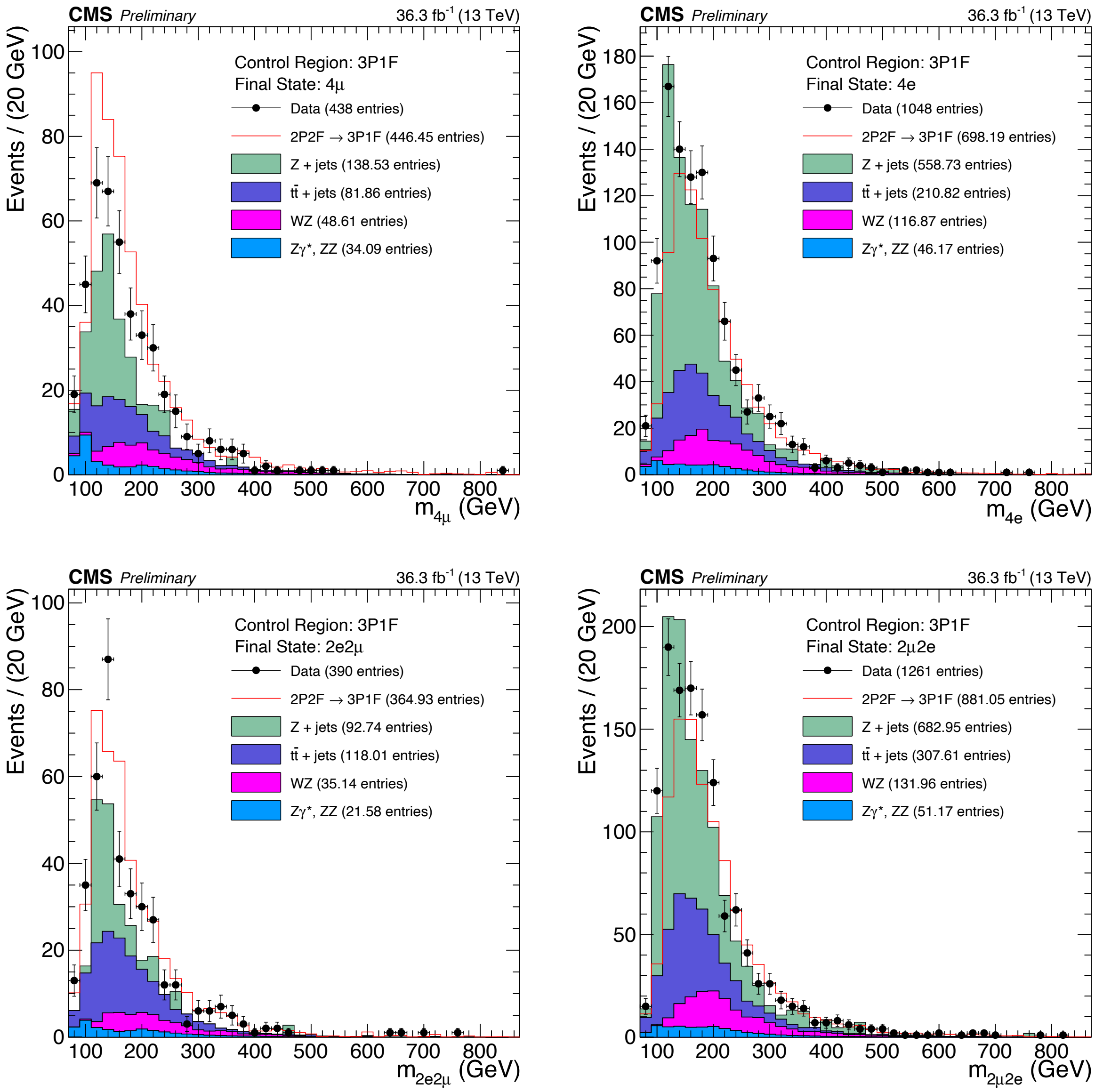

Figure 2-5. Events from 2016 UL data that pass 3P1F CR selections (black markers) are compared to the stacked 3P1F distributions of simulated samples (Z + jets, $t\bar{t}$ + jets, WZ, Z/γ^*, ZZ) and to the predicted contribution of 2-prompt-2-nonprompt-lepton processes to the 3P1F CR (red line). This prediction is obtained by making a distribution of all event weights given by Eq. 2-4 and stacking that on top of the WZ and ZZ distributions. The results are split into the 4ℓ final states: 4μ (top left), 4e (top right), $2e2\mu$ (bottom left), $2\mu 2e$ (bottom right).

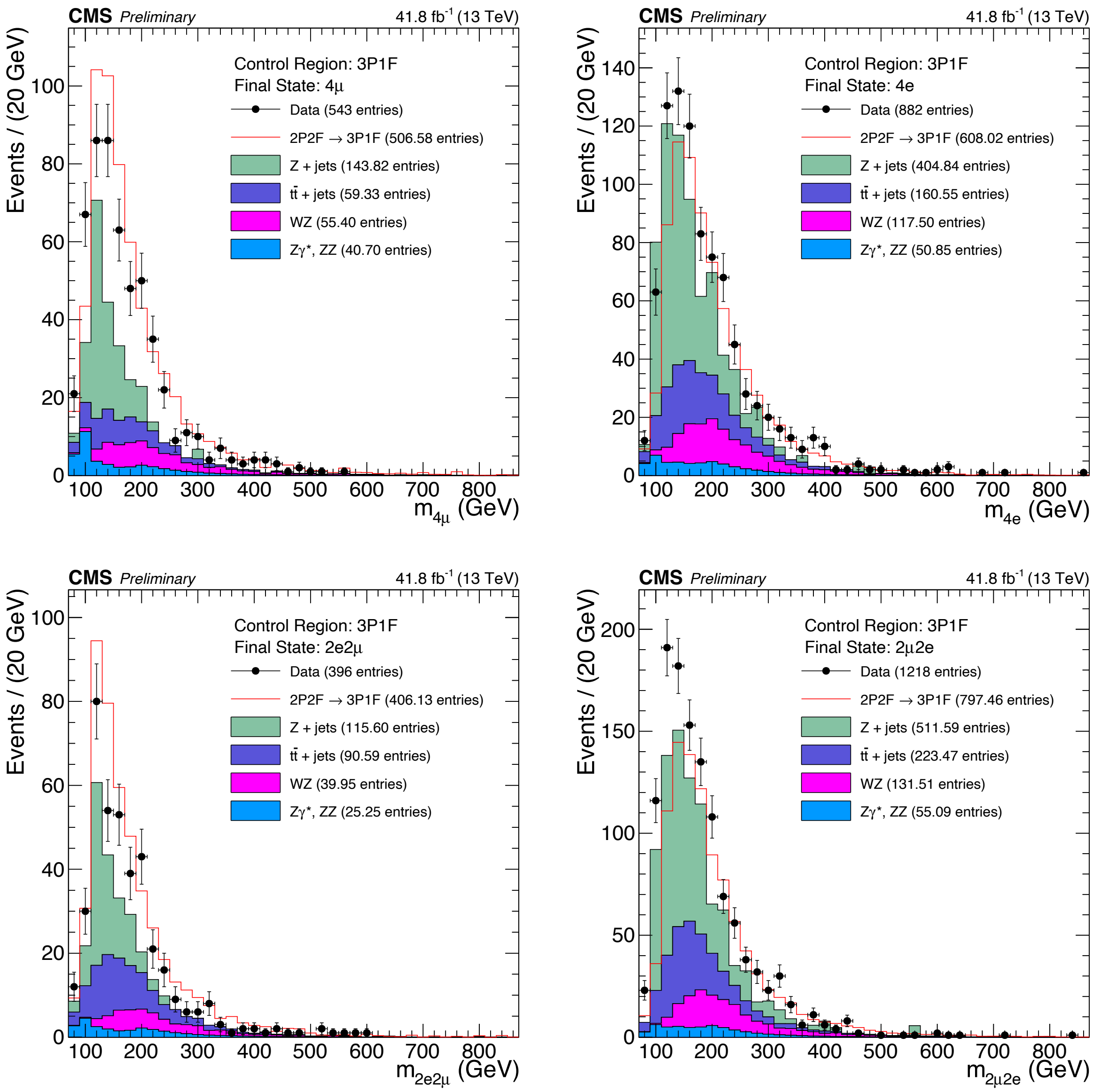

Figure 2-6. Events from 2017 UL data that pass 3P1F CR selections (black markers) are compared to the stacked 3P1F distributions of simulated samples (Z + jets, $t\bar{t}$ + jets, WZ, Z/γ^*, ZZ) and to the predicted contribution of 2-prompt-2-nonprompt-lepton processes to the 3P1F CR (red line). This prediction is obtained by making a distribution of all event weights given by Eq. 2-4 and stacking that on top of the WZ and ZZ distributions. The results are split into the 4ℓ final states: 4μ (top left), 4e (top right), $2e2\mu$ (bottom left), $2\mu 2e$ (bottom right).

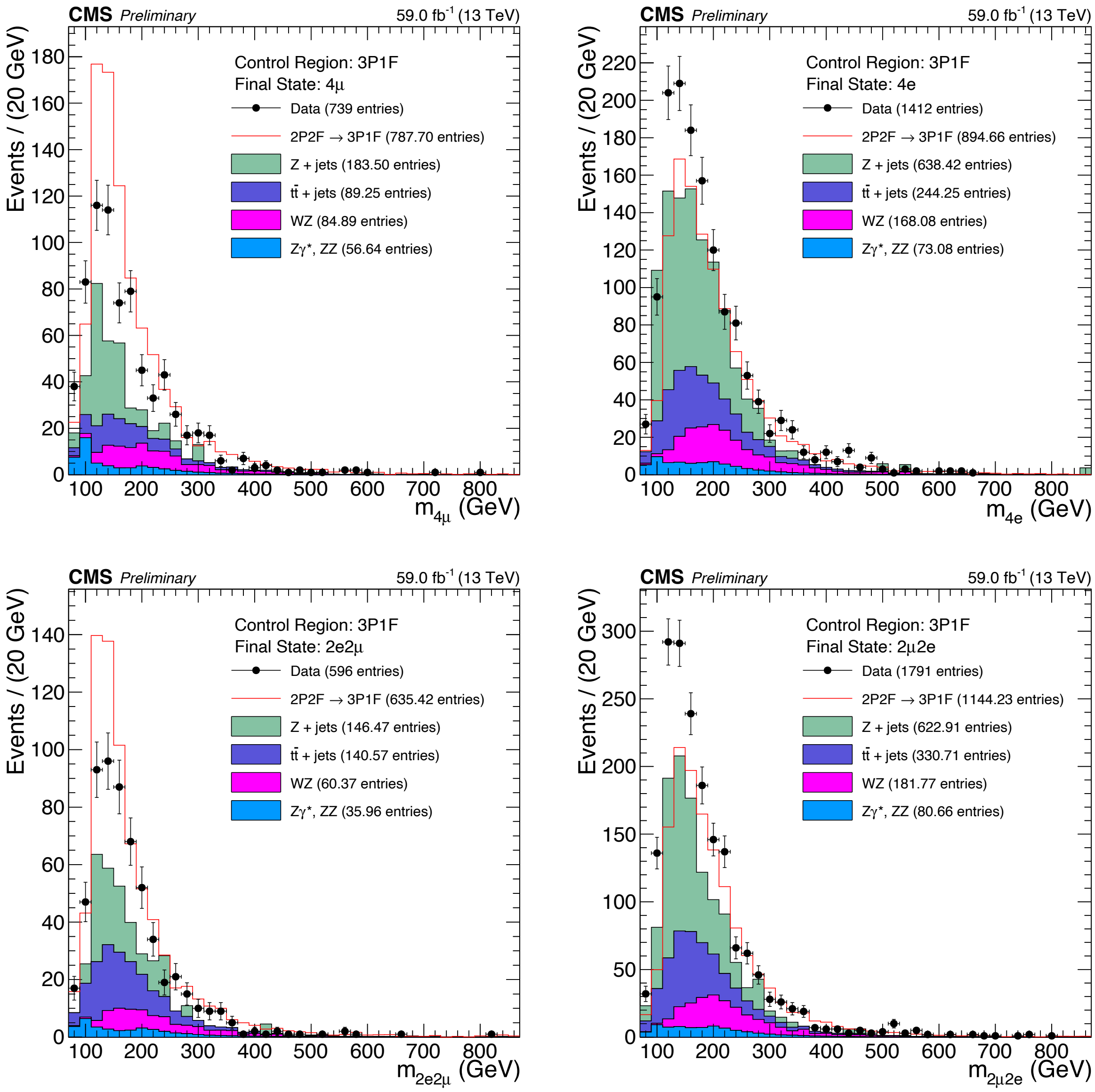

Figure 2-7. Events from 2018 UL data that pass 3P1F CR selections (black markers) are compared to the stacked 3P1F distributions of simulated samples (Z + jets, $t\bar{t}$ + jets, WZ, Z/γ^*, ZZ) and to the predicted contribution of 2-prompt-2-nonprompt-lepton processes to the 3P1F CR (red line). This prediction is obtained by making a distribution of all event weights given by Eq. 2-4 and stacking that on top of the WZ and ZZ distributions. The results are split into the 4ℓ final states: 4μ (top left), 4e (top right), 2e2μ (bottom left), 2μ2e (bottom right).

towards both the numerator and denominator of Eq. 2-14.

Events are selected that satisfy the following criteria:

- The event has exactly 3 leptons.
- The event contains $E_{\mathrm{T}}^{\mathrm{miss}} < 25$ GeV.
- Two of the leptons form a Z candidate. A Z candidate is formed when:
  - The lepton pair is OSSF.
  - Both leptons PTS.
  - The leading lepton has $p_{\mathrm{T}} > 20$ GeV.
  - The subleading lepton has $p_{\mathrm{T}} > 10$ GeV.
  - The lepton pair satisfies $\left|m_{\ell\ell} - m_{\mathrm{Z}_{\mathrm{PDG}}}\right| < 7$ GeV
- The third and final lepton is loose (and may be a PTS or FTS lepton).
- Suppress QCD processes: probe lepton and OS lepton from Z have $m_{\ell\ell} > 4$ GeV.

The calculation of f requires that ℓ_{L} is a nonprompt lepton but since data events were used, this may not be the case. For example, the decay of WZ produces 3 prompt leptons and so this contribution must be subtracted. Thus, the number of expected prompt probe leptons from WZ events is subtracted from both the numerator and denominator in Eq. 2-14 for each $\ell, p_{\mathrm{T}}, |\eta|$ bin. The final OS Method misidentification rates for electrons and muons are shown in Figure 2-8 using 2016–2018 UL data.

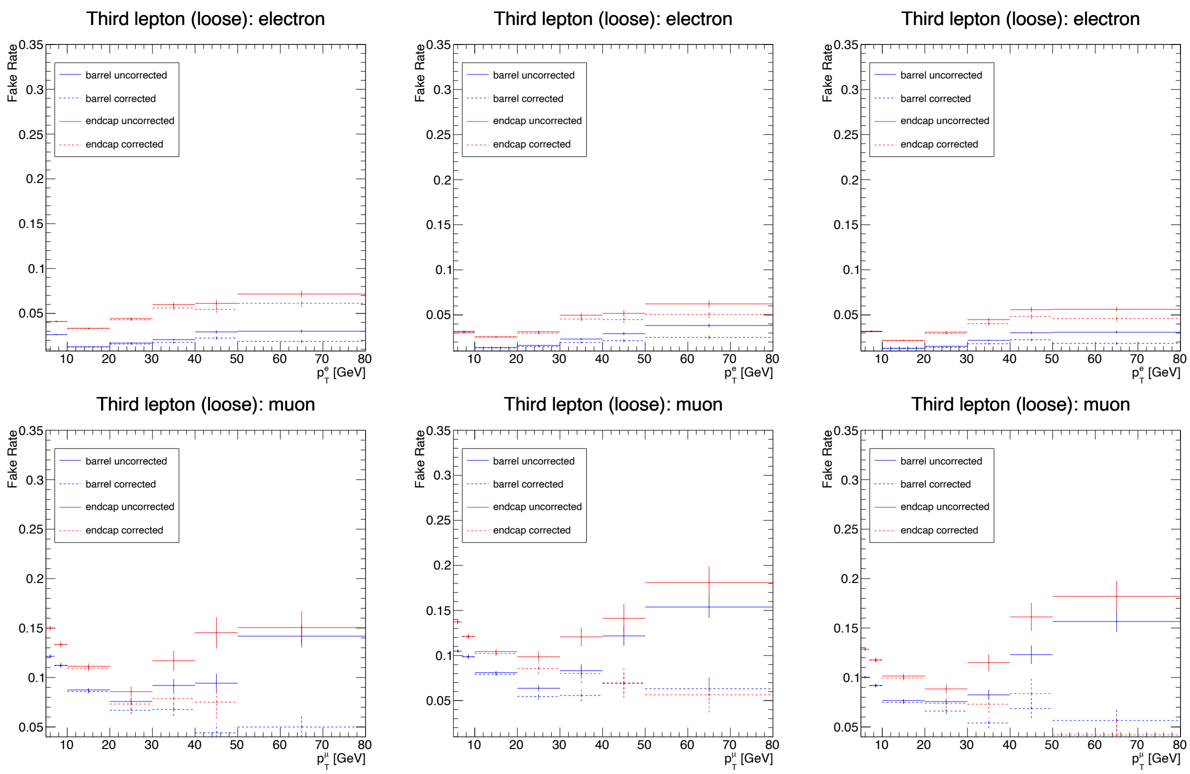

Figure 2-8. Electron (top row) and muon (bottom row) fake rates evaluated using the OS Method vs. the p_{T} of the probe lepton for 2016, 2017, and 2018 UL data (left, middle, right columns, respectively).

CHAPTER 3
SEARCH FOR LOW-MASS DILEPTON RESONANCES IN THE $H \rightarrow ZX/XX \rightarrow 4\ell$ CHANNEL

3.1 Introduction

Words.

words.

REFERENCES

[1] P.A. Zyla et al. Review of Particle Physics. *PTEP*, 2020(8):083C01, 2020. doi: 10.1093/ptep/ptaa104. and 2021 update.